Rotation Invariant Feature Extraction From 3-D Acceleration Signals

Takumi Kobayashi, Koiti Hasida and Nobuyuki Otsu

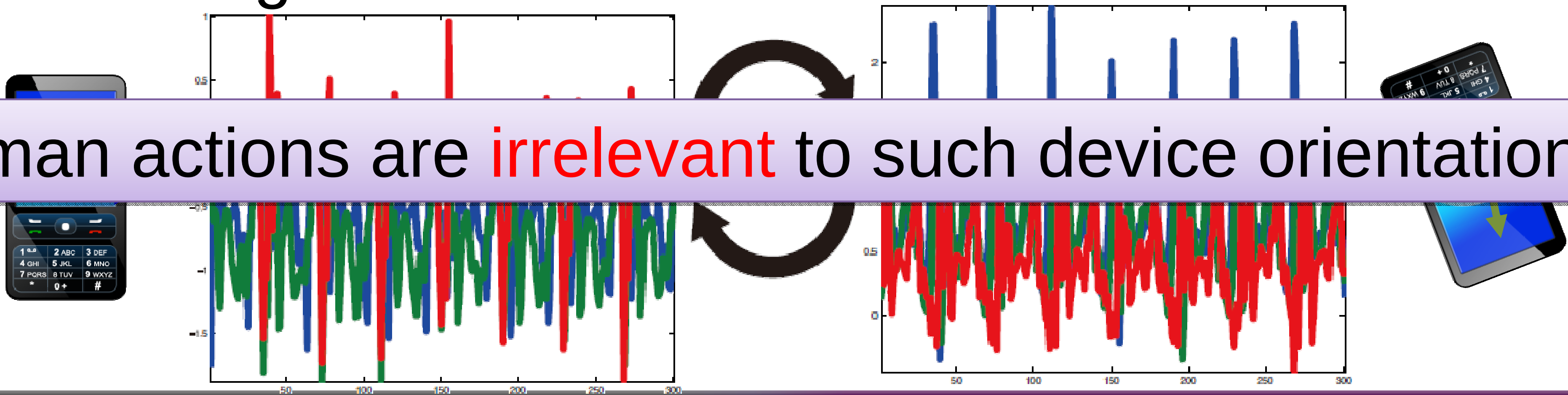
National Institute of Advanced Industrial Science and Technology (AIST), Japan

ICASSP 2011, Prague, Czech

Motivation

- 3-D accelerometer in smartphones has potential to provide various applications beyond simple screen-rotation.
  - Action recognition for identifying personal situation
  - Person identification such as by gait for security

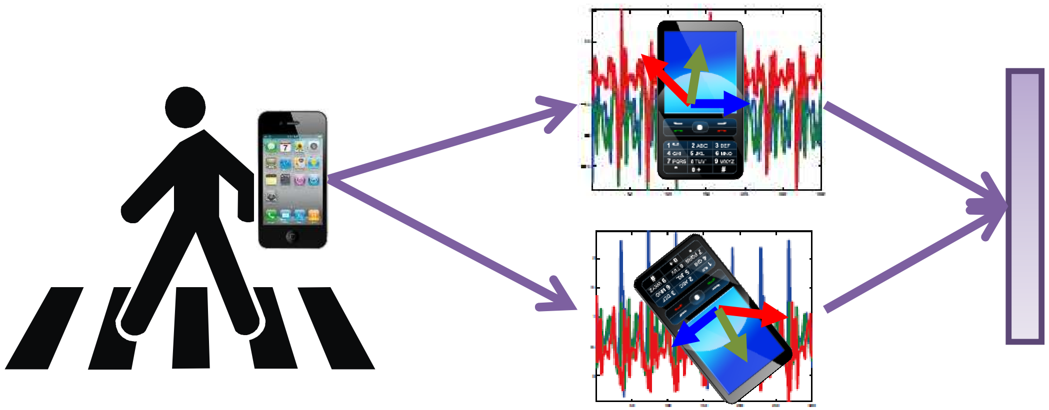
  etc..
- 3-D (x,y,z) signals by the accelerometer are rotated according to the device orientation.



Human actions are irrelevant to such device orientation.

Goal

- We propose an efficient method to extract rotation-invariant features from 3-D acceleration signals
  - without any preprocessing
  - for recognizing human actions independently of the device orientation.

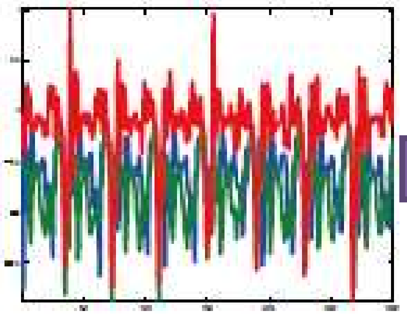


Rotation invariance:
The same feature (vector) no matter how the device is oriented.

- The proposed features are based on frequency characteristics by Fourier transform
  - considering correlations among all frequencies as well as ordinary power spectrum.

Related Works

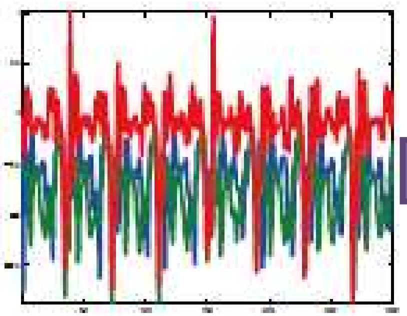
- Simple features (statistics) [Cho, et al., 2010]



Gravity vector $\boldsymbol{g}$ by moving average

$\left\{ \mu, \sigma, \text{min}, \text{max} \right\} \times \left\{ \text{L2-norm } \|\boldsymbol{a}_t\|, \text{ Dot-product } \boldsymbol{a}_t' \boldsymbol{g}, \text{ Cross-product } \boldsymbol{a}_t \times \boldsymbol{g} \right\}$

- Wavelet-based features [Mantyjarvi, et al., 2001]



PCA

$\left\{ 1^{st} \text{ PC}, 2^{nd} \text{ PC}, 3^{rd} \text{ PC} \right\} \times \left\{ \text{Several levels of Wavelets (Daubechies)} \right\}$

The statistical *preprocessing* compensates the signal orientations, though requiring plenty of signal data for stable results.

Proposed Method

- The frequency information would be useful to characterize human actions,
  - since the human actions are mainly consisting of periodic motions, such as walking, running and so on.

- (ST)FFT is applied to extract the frequency information from the input signals:

  $\boldsymbol{s}(t) \in \mathbb{R}^3$: 3-D signal

  $\boldsymbol{f}(\omega) = \int e^{-i\omega t}\,\boldsymbol{s}(t)dt \in \mathbb{C}^3$: 3-D Fourier features

  $\boldsymbol{F} = [\boldsymbol{f}(\omega_1), \ldots, \boldsymbol{f}(\omega_n)] \in \mathbb{C}^{3\times n}$: Fourier feature matrix

Proposed Method

- Correlation matrix of Fourier features $\boldsymbol{F}$ is given by

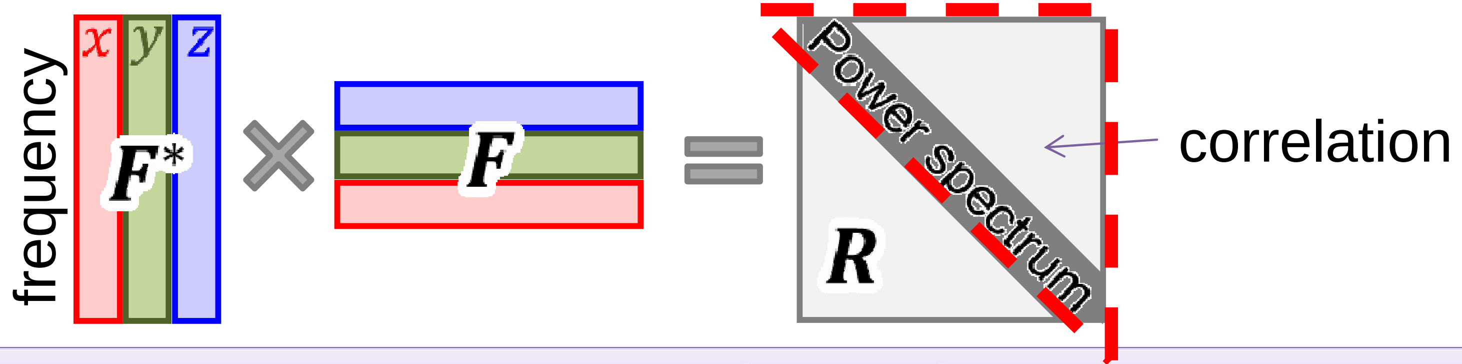
$$\boldsymbol{R} = \boldsymbol{F}^*\boldsymbol{F} \in \mathbb{C}^{n\times n}$$

  – Diagonal elements are ordinary power spectrum.

$$R_{jj} = \boldsymbol{f}(\omega_j)^*\boldsymbol{f}(\omega_j) = \sum_{d\in\{x,y,z\}}\left|f_d(\omega_j)\right|^2 \in \mathbb{R}$$

  – Off-diagonal elements are cross-correlation of frequencies.

$$R_{jk} = \boldsymbol{f}(\omega_j)^*\boldsymbol{f}(\omega_k) = \sum_{d\in\{x,y,z\}} f_d(\omega_j)^* f_d(\omega_k) \in \mathbb{C}$$



The proposed features are $\{|R_{jk}|\}_{j\le k}$,
absolute values of upper triangle components.

Invariance

- to Rotation

  $\boldsymbol{A} \in \mathbb{R}^{3\times 3}$: Rotation matrix, $\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{I}$ (unitary)

  $\hat{\boldsymbol{s}}(t) = \boldsymbol{A}\boldsymbol{s}(t)$: Rotated signal

  $\widehat{\boldsymbol{F}} = \boldsymbol{A}\boldsymbol{F}$: Rotated Fourier features

$$\boxed{\widehat{\boldsymbol{R}} = \widehat{\boldsymbol{F}}^*\widehat{\boldsymbol{F}} = \boldsymbol{F}^*\boldsymbol{A}'\boldsymbol{A}\boldsymbol{F} = \boldsymbol{F}^*\boldsymbol{F} = \boldsymbol{R}}$$

Rotation invariant

- to Temporal Shift (delay)

  $\tilde{\boldsymbol{s}}(t) = \boldsymbol{s}(t-\tau)$: Delayed signal

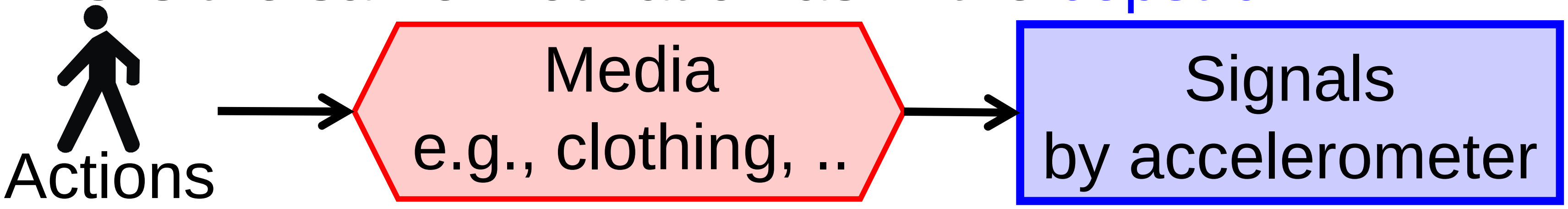
  $\tilde{\boldsymbol{f}}(\omega) = e^{-i\omega\tau}\boldsymbol{f}(\omega)$: Delayed Fourier features

$$\left|\tilde{R}_{jk}\right| = \left|\tilde{\boldsymbol{f}}(\omega_j)^*\tilde{\boldsymbol{f}}(\omega_k)\right| = \left|e^{i\omega_j\tau}e^{-i\omega_k\tau}\boldsymbol{f}(\omega_j)^*\boldsymbol{f}(\omega_k)\right|$$
$$= \left|\boldsymbol{f}(\omega_j)^*\boldsymbol{f}(\omega_k)\right| = \left|R_{jk}\right|$$

also, invariant

Post-processing

- Actual signals are transformed by a transfer function due to various media, such as closing.
- Thus, we take the logarithm of the extracted features to linearly separate those effects.
  - This is the same motivation as in the cepstrum.



$\boldsymbol{F}$ Transfer function $\boldsymbol{b} \in \mathbb{C}^{n\times 1}$ $\check{\boldsymbol{F}} = \boldsymbol{F}\,\mathrm{diag}(\boldsymbol{b})$

$$\log|\check{R}_{jk}| = \log|b_j^* R_{jk} b_k| = \log|R_{jk}| + \log|b_j^* b_k|$$

Linear multivariate analysis (FDA,etc) could statistically cancel out those biases (interferences).

Experiment

Task

Challenging human (gait) identification without any control for this experiment

- 58 persons freely walk in daily life with holding iPhone in hands.
- 2,331 signal sequences (avg. 40 seq. per person)

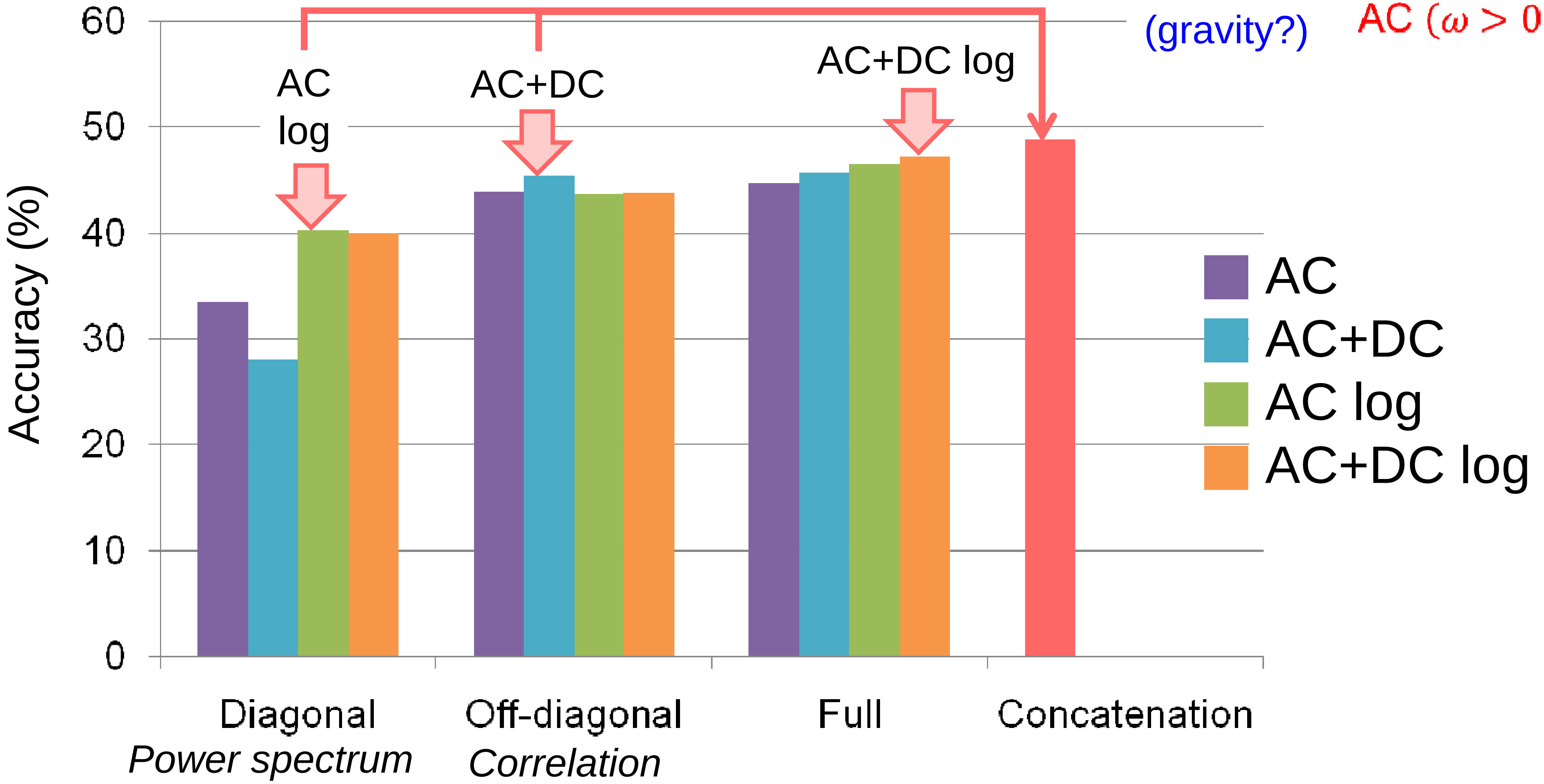
- Experimental setup
  - 70-frame running window for STFFT with 35 frames step size.
  - Multi-class classification by nearest means in Fisher discriminant space and by majority voting.
  - We measured averaged accuracy across persons by using 3-fold cross validation.

Experimental Results (1)

- Various settings in our method

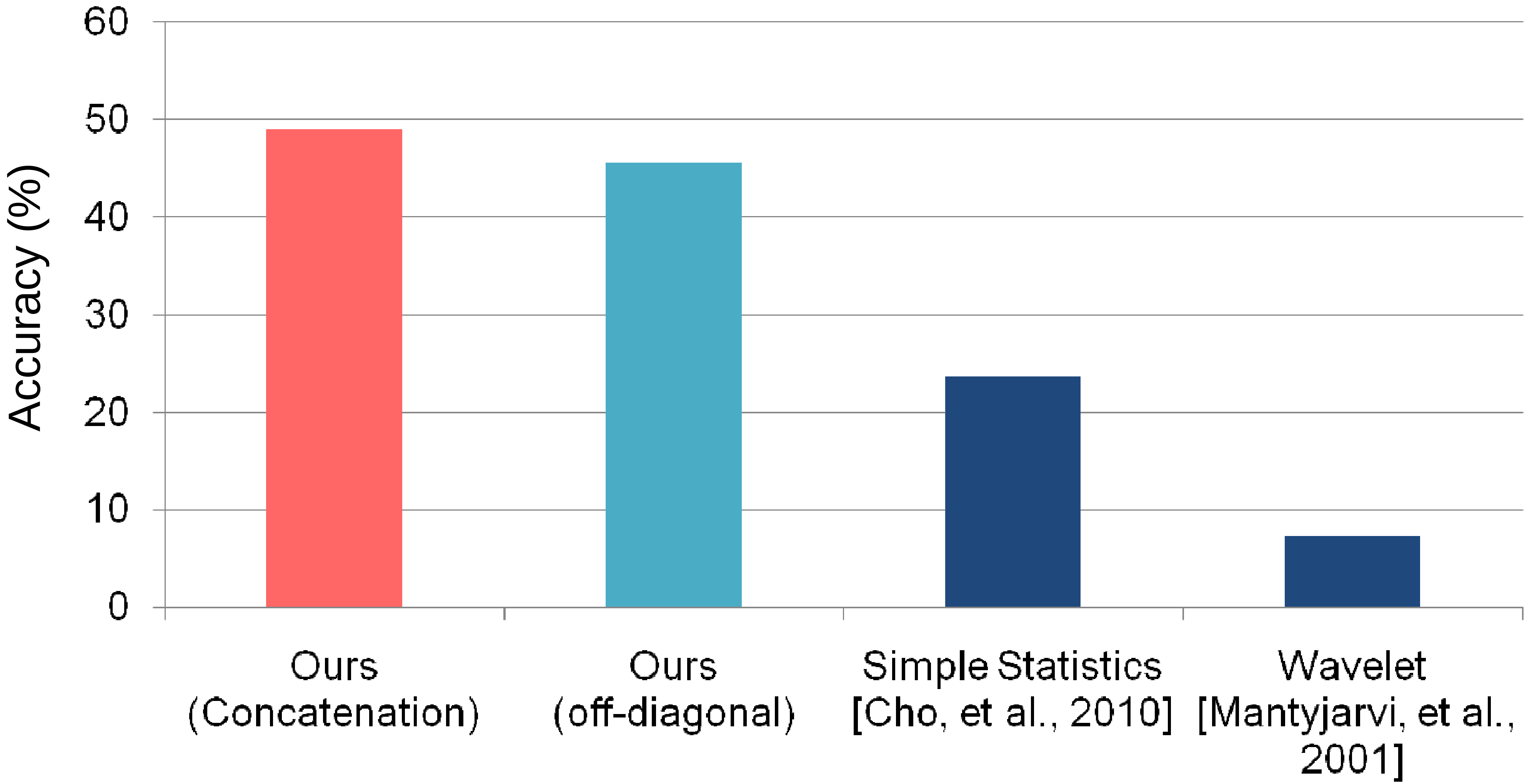
Fourier features

$$\boldsymbol{F} = [\underbrace{\boldsymbol{f}(0)}_{\text{DC (gravity?)}}, \underbrace{\ldots, \boldsymbol{f}(\omega_n)}_{\text{AC } (\omega > 0)}]$$



Experimental Results (2)

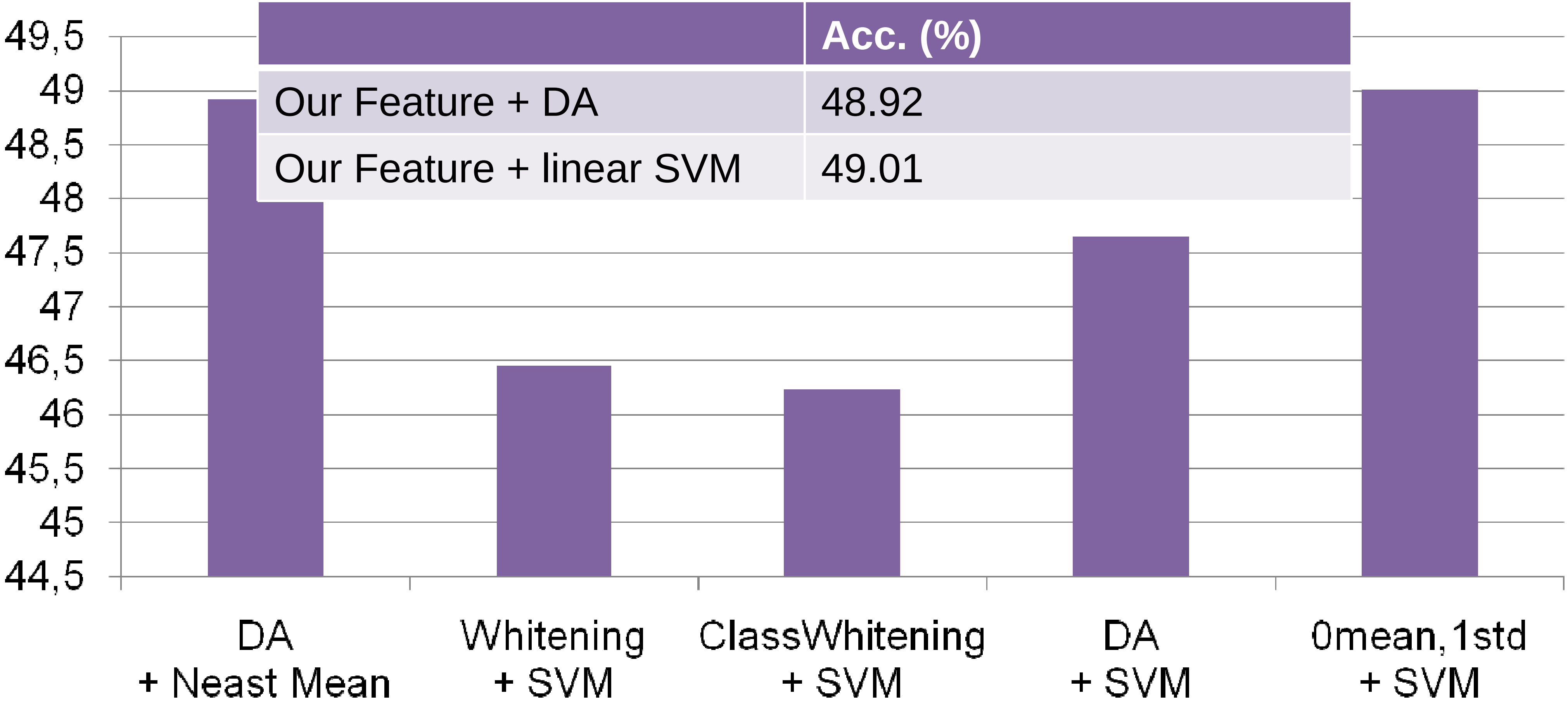
- Comparison to the others



Experimental Results (supplemental)

- SVM classifier

	Acc. (%)
Our Feature + DA	48.92
Our Feature + linear SVM	49.01



Conclusion

- We have proposed the method to extract rotation-invariant features from 3D acceleration signals.
  - Based on the correlation matrix of Fourier frequency features, containing both power spectrum (diagonal) and cross-correlations (off-diagonal) of frequencies.
  - Inherently invariant to rotations as well as temporal shift without any ad-hoc preprocessing.
- In the experiment on human identification, the proposed method produced favorable performances compared to the others.
- The method is so general as to be applied to the other tasks using acceleration, e.g., segmenting and recognizing various human activities.

Thank you for your attentions!!